Many formal models of abstraction in language exist, but here we focus on work describing prediction and the emergence of abstract category structure as a function of accumulating knowledge of the contexts in which experience is grounded (Elman 1990; see also Altmann 1997). Jeff Elman’s work with the simple recurrent network (SRN) is the quintessential example from a computational standpoint of abstraction, prediction, and the relationship between the two (Elman 1990; 1993). Through accumulated experience of sequences of words, categorical distinctions such as between parts of speech (e.g., noun and verb), and between classes of nouns and verbs (e.g., edible objects and intransitive verbs) emerge in a network given the task of predicting the next word in the sequence.

Gilead et al. perceive an insufficiency in models exhibiting an “undifferentiated, continuous hierarchy of mental representations of different levels of abstractness.” Yet Elman’s SRN was undifferentiated computationally (hidden layer units all functioned identically). After learning though, it was not undifferentiated functionally. Similarity relationships in its equivalent of the external world (language input to the SRN) were maintained in its acquired internal representations, and these allowed the SRN to predict the space of possible inputs at the next point in time. Hierarchy was only categorical to the extent that hierarchical clustering is categorical (different clusters would exhibit different hierarchies). Abstraction in Elman’s work was guided, meaning generalization was graded also – a desirable property in a probabilistic world. Importantly, and unlike Gilead et al.’s framework, the computational principles that underpin the successes of the SRN (which have since been shown, in deep recurrent neural networks to scale up to the demands of realistically large vocabularies) are general principles of learning and development (Elman et al. 1996).

The emergence of increasingly abstract representations (not just in language) may rely on domain-general neurological mechanisms for tracking systematics across space and time (for discussion of how one such mechanism may apply to abstract concepts, see Davis et al. 2020). However, a problem for any experience-based model of abstraction is how we sample enough of the world to track those systematics and converge on shared meaning. Here, language comes in again: It allows us to experience more of the world than we could via direct experience alone. Experiencing spoken, signed, and written words – and their distributional patterns of co-occurrence both with other words in sentences and with the real world – opens a window into other people’s (embodied) experiences. Distributional language statistics are a rich source of knowledge (e.g., Louwiser 2008), enabling us to make predictions about things not directly experienced.

Language also facilitates prediction and abstraction in ways non-linguistic thought does not. For example, labels may penetrate through the representational gradient by operating directly on mental states (Elman 2009). Although classical thinking holds that language is merely a means to communicating our thoughts, more recent work has shown that language has a functional role not only in higher-order thought, but also perception (for a review, see Lupyan 2012). A consequence of language’s influence across the gradient of abstraction is that concepts do not operate only at the modality-specific level: labels may (among other things) help integrate modality specific information
in higher-order association areas. Gilead et al. cite meta-analytic findings that lexical-semantic tasks tend to activate higher-order brain regions far removed from modality-specific areas (Binder et al. 2009) as “compelling evidence” against distributed, modality-specific models of cognition. But, these activated higher-order regions are integral to multimodal integration and conceptual access via labels. Furthermore, because there is diversity in the modalities in which different things are experienced (e.g., sunsets visually, vs. thunder auditorily), conceptual representations reflect that diversity (e.g., Davis et al. in press). Thus, when experiments average over dozens of diverse concepts, activity in the various modalities that contribute to each one is likely to be washed out.

Abstraction is a process, and this process engenders a gradient, not qualitatively distinct levels in a representational hierarchy. Moreover, an account emphasizing “representational diversity” to address how humans use prediction and abstraction to transcend the present moment should recognize the ubiquitous role of language. Not only does the scientific study of language processing offer well-tested, formalized frameworks for understanding how abstract structure emerges (e.g., Elman 1990; see also Altmann 2017), but language itself plays a functional role in facilitating “mental travel” via its integral role in prediction and abstraction.


Davis, C. P., Joergensen, G. H., Boddy, P., Dowling, C., & Yee, E. (in press) Making it harder to “see” meaning: The more you see something, the more its conceptual representation is susceptible to visual interference. Psychological Science. [CPD]


